Cassandra Crossing/ IA: il Collasso, reloaded

(575)—Perché un dettaglio tecnico delle false IA, per giunta di tipo matematico, dovrebbe essere conosciuto da tutti?

Cassandra Crossing/ IA: il Collasso, reloaded

Figure 1:

(575)—Perché un dettaglio tecnico delle false IA, per giunta di tipo matematico, dovrebbe essere conosciuto da tutti?

29 febbraio 2024—Chi dei 24 indefettibili lettori si fosse già cimentato nella lettura della precedente esternazione di Cassandra sullo stesso tema, può tranquillamente proseguire; **tutti gli altri dovrebbero proprio leggerla**, come se fosse il primo tempo di un film giallo a cui siete arrivati in ritardo.

Inizia il secondo tempo, il cui sceneggiatore non è Cassandra, ma una meritoria persona che ha postato su Reddit un riassunto ben fatto di una abbastanza astrusa questione, descritta in una paper specialistica.

Disclosure: Cassandra l'ha rielaborato, tradotto (con un LLM) e poi nuovamente rielaborato, e qui lo inserisce. Un grazie d'obbligo al vero autore, quindi, e procediamo.

Warning: un minimo di matematica è necessaria, ma se avete mai fatto un esamino od una tesina di statistica base, la cosa sarà per voi chiarissima, anzi abbagliante nella sua semplicità.

Il *collasso di un modell*o di LLM avviene quando un modello, addestrato sui dati generati da generazioni precedenti di modelli, inizia a perdere informazioni, particolarmente sulle code della distribuzione statistica dei dati originali, ed alla fine converge verso una stima a punto singolo, con poca varianza statistica. (*N.d.R: in soldoni, fornisce sempre la stessa risposta a qualsiasi domanda*).

Il *collasso del modello* si verifica a causa di due fonti di errore: errore di approssimazione statistica dovuto al campionamento finito, ed errore di approssimazione funzionale dovuto a modelli imperfetti. Questi errori si accumulano nel corso delle generazioni, facendo sì che la distribuzione stimata si allontani ulteriormente dall'originale.

Alcuni studi mostrano che il collasso del modello si verifica in modelli semplici come i *Gaussian Mixture Models* ed i *Variational Autoencoders*, nonché in *Modelli Linguistici* più complessi. Anche una taratura del modello durante l'addestramento non impedisce il *collasso del modello* nei modelli linguistici.

Nel corso del tempo, i dati generati dai modelli affetti da *collasso del modello* iniziano a contenere sequenze improbabili, e perdono informazioni sulle code statistiche della distribuzione originale.

I ricercatori sostengono che, per evitare il collasso del modello e poter mantenere le sue prestazioni, l'accesso a dati generati direttamente dagli esseri umani, da usare per il training, rimarrà essenziale. Inoltre, i dati prodotti durante le interazioni umane con i modelli linguistici saranno sempre più preziosi.

In sintesi, si evidenzia un fenomeno importante in cui i modelli addestrati ricorsivamente sui propri dati generati iniziano a perdere fedeltà, ed alla fine convergono verso uno stato non ottimale.

L'accesso a dati originali generati dall'uomo è quindi necessario per evitare il collasso del modello, e sostenerne le prestazioni nel lungo periodo. Per questo motivo, man mano che i modelli linguistici inizieranno a generare una crescente quantità dei contenuti del web, distinguere su larga scala i dati umani da quelli generati dai modelli diventerà una sfida importante per riuscire ad allenare correttamente i modelli linguistici.

Cassandra, facendo stavolta davvero il proprio mestiere, termina con un avvertimento.

Sì, parliamo proprio delle *false intelligenze artificiali* che trovate già adesso nelle ultime versioni dei prodotti che usate. Persino nei generatori di codici informatici, tipo Copilot, che la maggior parte degli sviluppatori ormai usano. E quindi i cui effetti trovate nel software che usate e *subite* tutti i giorni, e che controllerà la prossima versione degli *oggetti intelligenti*.

Stateve accuorti!

Scrivere a Cassandra—Twitter—Mastodon
Videorubrica "Quattro chiacchiere con Cassandra"

Lo Slog (Static Blog) di Cassandra
L'archivio di Cassandra: scuola, formazione e pensiero

By Marco A. L. Calamari on February 29, 2024.

Canonical link

Exported from Medium on August 27, 2025.